0120   OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/940,166

DATE: 10/25/2001
TIME: 10:19:05

Input Set : N:\Crf3\RULE60\09940166.txt
Output Set: N:\CRF3\10252001\I940166.raw

SEQUENCE LISTING

```
  3 (1) GENERAL INFORMATION:
  5      (i) APPLICANT: Blank, Gregory S.
  6                     Narindray, Daljit S.
  7                     Zapata, Gerardo A.
  9     (ii) TITLE OF INVENTION: Protein Recovery
 11    (iii) NUMBER OF SEQUENCES: 7
 13     (iv) CORRESPONDENCE ADDRESS:
 14          (A) ADDRESSEE: Genentech, Inc.
 15          (B) STREET: 1 DNA Way
 16          (C) CITY: South San Francisco
 17          (D) STATE: California
 18          (E) COUNTRY: USA
 19          (F) ZIP: 94080
 21      (v) COMPUTER READABLE FORM:
 22          (A) MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 23          (B) COMPUTER: IBM PC compatible
 24          (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 25          (D) SOFTWARE: WinPatin (Genentech)
 27     (vi) CURRENT APPLICATION DATA:
 28          (A) APPLICATION NUMBER: US/09/940,166
 29          (B) FILING DATE: 27-Aug-2001
 30          (C) CLASSIFICATION:
 32    (vii) PRIOR APPLICATION DATA:
 33          (A) APPLICATION NUMBER: 09/097,309
 34          (B) FILING DATE: 13-JUN-1997
 36   (viii) ATTORNEY/AGENT INFORMATION:
 37          (A) NAME: Schwartz, Timothy R.
 38          (B) REGISTRATION NUMBER: 32171
 39          (C) REFERENCE/DOCKET NUMBER: P1105R1
 41     (ix) TELECOMMUNICATION INFORMATION:
 42          (A) TELEPHONE: 650/225-7467
 43          (B) TELEFAX: 650/952-9881
 44 (2) INFORMATION FOR SEQ ID NO: 1:
 46      (i) SEQUENCE CHARACTERISTICS:
 47          (A) LENGTH: 241 amino acids
 48          (B) TYPE: Amino Acid
 49          (D) TOPOLOGY: Linear
 51     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 53 Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly
 54  1               5                  10                  15
 56 Gly Ser Leu Arg Leu Ser Cys Ala Thr Ser Gly Tyr Thr Phe Thr
 57                 20                  25                  30
 59 Glu Tyr Thr Met His Trp Met Arg Gln Ala Pro Gly Lys Gly Leu
 60                 35                  40                  45
 62 Glu Trp Val Ala Gly Ile Asn Pro Lys Asn Gly Gly Thr Ser His
 63                 50                  55                  60
```

C--> 28
C--> 29

ENTERED

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/940,166

Input Set : N:\Crf3\RULE60\09940166.txt
Output Set: N:\CRF3\10252001\I940166.raw

```
 65    Asn Gln Arg Phe Met Asp Arg Phe Thr Ile Ser Val Asp Lys Ser
 66                         65                  70                  75
 68    Thr Ser Thr Ala Tyr Met Gln Met Asn Ser Leu Arg Ala Glu Asp
 69                         80                  85                  90
 71    Thr Ala Val Tyr Tyr Cys Ala Arg Trp Arg Gly Leu Asn Tyr Gly
 72                         95                  100                 105
 74    Phe Asp Val Arg Tyr Phe Asp Val Trp Gly Gln Gly Thr Leu Val
 75                         110                 115                 120
 77    Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu
 78                         125                 130                 135
 80    Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly
 81                         140                 145                 150
 83    Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp
 84                         155                 160                 165
 86    Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val
 87                         170                 175                 180
 89    Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val
 90                         185                 190                 195
 92    Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val Asn
 93                         200                 205                 210
 95    His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys
 96                         215                 220                 225
 98    Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu
 99                         230                 235                 240
101    Leu
102    241
104 (2) INFORMATION FOR SEQ ID NO: 2:
106        (i) SEQUENCE CHARACTERISTICS:
107            (A) LENGTH: 214 amino acids
108            (B) TYPE: Amino Acid
109            (D) TOPOLOGY: Linear
111        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
113    Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val
114     1               5                  10                  15
116    Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Asp Ile Asn
117                     20                  25                  30
119    Asn Tyr Leu Asn Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys
120                     35                  40                  45
122    Leu Leu Ile Tyr Tyr Thr Ser Thr Leu His Ser Gly Val Pro Ser
123                     50                  55                  60
125    Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Thr Leu Thr Ile
126                     65                  70                  75
128    Ser Ser Leu Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln
129                     80                  85                  90
131    Gly Asn Thr Leu Pro Pro Thr Phe Gly Gln Gly Thr Lys Val Glu
132                     95                  100                 105
134    Ile Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro
135                     110                 115                 120
137    Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/940,166

DATE: 10/25/2001
TIME: 10:19:05

Input Set : N:\Crf3\RULE60\09940166.txt
Output Set: N:\CRF3\10252001\I940166.raw

```
138                        125              130                   135
140   Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val
141                        140              145                   150
143   Asp Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu
144                        155              160                   165
146   Gln Asp Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr
147                        170              175                   180
149   Leu Ser Lys Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu
150                        185              190                   195
152   Val Thr His Gln Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn
153                        200              205                   210
155   Arg Gly Glu Cys
156                        214
158 (2) INFORMATION FOR SEQ ID NO: 3:
160        (i) SEQUENCE CHARACTERISTICS:
161              (A) LENGTH: 36 amino acids
162              (B) TYPE: Amino Acid
163              (D) TOPOLOGY: Linear
165        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
167   Leu Gly Gly Arg Met Lys Gln Leu Glu Asp Lys Val Glu Glu Leu
168    1               5                  10                      15
170   Leu Ser Lys Asn Tyr His Leu Glu Asn Glu Val Ala Arg Leu Lys
171                   20                  25                      30
173   Lys Leu Val Gly Glu Arg
174                   35  36
176 (2) INFORMATION FOR SEQ ID NO: 4:
178        (i) SEQUENCE CHARACTERISTICS:
179              (A) LENGTH: 7 amino acids
180              (B) TYPE: Amino Acid
181              (D) TOPOLOGY: Linear
183        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
W--> 185  Leu Xaa Xaa Xaa Xaa Xaa Xaa
186    1               5       7
188 (2) INFORMATION FOR SEQ ID NO: 5:
190        (i) SEQUENCE CHARACTERISTICS:
191              (A) LENGTH: 2143 base pairs
192              (B) TYPE: Nucleic Acid
193              (C) STRANDEDNESS: Single
194              (D) TOPOLOGY: Linear
196        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
199   GAATTCAACT TCTCCATACT TTGGATAAGG AAATACAGAC ATGAAAAATC 50
201   TCATTGCTGA GTTGTTATTT AAGCTTTGGA GATTATCGTC ACTGCAATGC 100
203   TTCGCAATAT GGCGCAAAAT GACCAACAGC GGTTGATTGA TCAGGTAGAG 150
205   GGGGCGCTGT ACGAGGTAAA GCCCGATGCC AGCATTCCTG ACGACGATAC 200
207   GGAGCTGCTG CGCGATTACG TAAAGAAGTT ATTGAAGCAT CCTCGTCAGT 250
209   AAAAAGTTAA TCTTTTCAAC AGCTGTCATA AAGTTGTCAC GGCCGAGACT 300
211   TATAGTCGCT TTGTTTTTAT TTTTTAATGT ATTTGTAACT AGAATTCGAG 350
213   CTCGCCGGGG ATCCTCTAGA GGTTGAGGTG ATTTTATGAA AAAGAATATC 400
215   GCATTTCTTC TTGCATCTAT GTTCGTTTTT TCTATTGCTA CAAACGCGTA 450
```

```
217  CGCTGATATC CAGATGACCC AGTCCCCGAG CTCCCTGTCC GCCTCTGTGG 500
219  GCGATAGGGT CACCATCACC TGTCGTGCCA GTCAGGACAT CAACAATTAT 550
221  CTGAACTGGT ATCAACAGAA ACCAGGAAAA GCTCCGAAAC TACTGATTTA 600
223  CTATACCTCC ACCCTCCACT CTGGAGTCCC TTCTCGCTTC TCTGGTTCTG 650
225  GTTCTGGGAC GGATTACACT CTGACCATCA GCAGTCTGCA ACCGGAGGAC 700
227  TTCGCAACTT ATTACTGTCA GCAAGGTAAT ACTCTGCCGC CGACGTTCGG 750
229  ACAGGGCACG AAGGTGGAGA TCAAACGAAC TGTGGCTGCA CCATCTGTCT 800
231  TCATCTTCCC GCCATCTGAT GAGCAGTTGA AATCTGGAAC TGCCTCTGTT 850
233  GTGTGCCTGC TGAATAACTT CTATCCCAGA GAGGCCAAAG TACAGTGGAA 900
235  GGTGGATAAC GCCCTCCAAT CGGGTAACTC CCAGGAGAGT GTCACAGAGC 950
237  AGGACAGCAA GGACAGCACC TACAGCCTCA GCAGCACCCT GACGCTGAGC 1000
239  AAAGCAGACT ACGAGAAACA CAAAGTCTAC GCCTGCGAAG TCACCCATCA 1050
241  GGGCCTGAGC TCGCCCGTCA CAAAGAGCTT CAACAGGGGA GAGTGTTAAG 1100
243  CTGATCCTCT ACGCCGGACG CATCGTGGCG CTAGTACGCA AGTTCACGTA 1150
245  AAAACGGTAT CTAGAGGTTG AGGTGATTTT ATGAAAAAGA ATATCGCATT 1200
247  TCTTCTTGCA TCTATGTTCG TTTTTTCTAT TGCTACAAAC GCGTACGCTG 1250
249  AGGTTCAGCT GGTGGAGTCT GGCGGTGGCC TGGTGCAGCC AGGGGGCTCA 1300
251  CTCCGTTTGT CCTGTGCAAC TTCTGGCTAC ACCTTTACCG AATACACTAT 1350
253  GCACTGGATG CGTCAGGCCC CGGGTAAGGG CCTGGAATGG GTTGCAGGGA 1400
255  TTAATCCTAA AAACGGTGGT ACCAGCCACA ACCAGAGGTT CATGGACCGT 1450
257  TTCACTATAA GCGTAGATAA ATCCACCAGT ACAGCCTACA TGCAAATGAA 1500
259  CAGCCTGCGT GCTGAGGACA CTGCCGTCTA TTATTGTGCT AGATGGCGAG 1550
261  GCCTGAACTA CGGCTTTGAC GTCCGTTATT TTGACGTCTG GGGTCAAGGA 1600
263  ACCCTGGTCA CCGTCTCCTC GGCCTCCACC AAGGGCCCAT CGGTCTTCCC 1650
265  CCTGGCACCC TCCTCCAAGA GCACCTCTGG GGGCACAGCG GCCCTGGGCT 1700
267  GCCTGGTCAA GGACTACTTC CCCGAACCGG TGACGGTGTC GTGGAACTCA 1750
269  GGCGCCCTGA CCAGCGGCGT GCACACCTTC CCGGCTGTCC TACAGTCCTC 1800
271  AGGACTCTAC TCCCTCAGCA GCGTGGTGAC CGTGCCCTCC AGCAGCTTGG 1850
273  GCACCCAGAC CTACATCTGC AACGTGAATC ACAAGCCCAG CAACACCAAG 1900
275  GTCGACAAGA AAGTTGAGCC CAAATCTTGT GACAAAACTC ACACATGCCC 1950
277  GCCGTGCCCA GCACCAGAAC TGCTGGGCGG CCGCATGAAA CAGCTAGAGG 2000
279  ACAAGGTCGA AGAGCTACTC TCCAAGAACT ACCACCTAGA GAATGAAGTG 2050
281  GCAAGACTCA AAAAGCTTGT CGGGGAGCGC TAAGCATGCG ACGGCCCTAG 2100
283  AGTCCCTAAC GCTCGGTTGC CGCCGGGCGT TTTTTATTGT TAA 2143
285  (2) INFORMATION FOR SEQ ID NO: 6:
287       (i) SEQUENCE CHARACTERISTICS:
288          (A) LENGTH: 237 amino acids
289          (B) TYPE: Amino Acid
290          (D) TOPOLOGY: Linear
292       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
294  Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe
295  -23       -20              -15                 -10
297  Ser Ile Ala Thr Asn Ala Tyr Ala Asp Ile Gln Met Thr Gln Ser
298           -5                    1               5
300  Pro Ser Ser Leu Ser Ala Ser Val Gly Asp Arg Val Thr Ile Thr
301           10                   15                  20
303  Cys Arg Ala Ser Gln Asp Ile Asn Asn Tyr Leu Asn Trp Tyr Gln
304      25                       30                  35
306  Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr Tyr Thr Ser
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/940,166

Input Set : N:\Crf3\RULE60\09940166.txt
Output Set: N:\CRF3\10252001\I940166.raw

```
                           45                    50
307            40
309  Thr Leu His Ser Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser
310            55                    60                    65
312  Gly Thr Asp Tyr Thr Leu Thr Ile Ser Ser Leu Gln Pro Glu Asp
313            70                    75                    80
315  Phe Ala Thr Tyr Tyr Cys Gln Gln Gly Asn Thr Leu Pro Pro Thr
316            85                    90                    95
318  Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr Val Ala Ala
319            100                   105                   110
321  Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu Lys Ser
322            115                   120                   125
324  Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro Arg
325            130                   135                   140
327  Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly
328            145                   150                   155
330  Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
331            160                   165                   170
333  Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu
334            175                   180                   185
336  Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser
337            190                   195                   200
339  Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
340            205                   210                   214
342  (2) INFORMATION FOR SEQ ID NO: 7:
344         (i) SEQUENCE CHARACTERISTICS:
345             (A) LENGTH: 300 amino acids
346             (B) TYPE: Amino Acid
347             (D) TOPOLOGY: Linear
349         (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
351  Met Lys Lys Asn Ile Ala Phe Leu Leu Ala Ser Met Phe Val Phe
352  -23            -20                   -15                   -10
354  Ser Ile Ala Thr Asn Ala Tyr Ala Glu Val Gln Leu Val Glu Ser
355                   -5                    1                     5
357  Gly Gly Gly Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys
358            10                    15                    20
360  Ala Thr Ser Gly Tyr Thr Phe Thr Glu Tyr Thr Met His Trp Met
361            25                    30                    35
363  Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Ala Gly Ile Asn
364            40                    45                    50
366  Pro Lys Asn Gly Gly Thr Ser His Asn Gln Arg Phe Met Asp Arg
367            55                    60                    65
369  Phe Thr Ile Ser Val Asp Lys Ser Thr Ser Thr Ala Tyr Met Gln
370            70                    75                    80
372  Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala
373            85                    90                    95
375  Arg Trp Arg Gly Leu Asn Tyr Gly Phe Asp Val Arg Tyr Phe Asp
376            100                   105                   110
378  Val Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ala Ser Thr
379            115                   120                   125
```

10/25/01

VERIFICATION SUMMARY
PATENT APPLICATION: **US/09/940,166**

DATE: 10/25/2001
TIME: 10:19:06

Input Set : **N:\Crf3\RULE60\09940166.txt**
Output Set: **N:\CRF3\10252001\I940166.raw**

L:28 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:29 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:185 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4